

# Xinya Du

---

CONTACT	Assistant Professor Department of Computer Science The University of Texas at Dallas 800 West Campbell Road, Richardson, TX 75080	Email: <a href="mailto:xinya.du@utdallas.edu">xinya.du@utdallas.edu</a> Website: <a href="https://xinyadu.github.io">https://xinyadu.github.io</a> Google Scholar
RESEARCH INTERESTS	Natural Language Processing, Computational Linguistics, Machine Learning, Deep Learning	
EDUCATION	<b>Cornell University</b> <i>Ph.D. in Computer Science (M.S. degree granted in Aug 2019)</i> Dissertation: Towards More Intelligent Extraction of Information from Documents Advisor: Claire Cardie	Aug 2016 – Aug 2021
	<b>Shanghai Jiao Tong University</b> <i>B.E. in Computer Science and Engineering</i>	Sep 2012 – Aug 2016
EXPERIENCE	<b>University of Texas at Dallas, Richardson, TX</b> <i>Assistant Professor in Computer Science</i>	Aug 2022 – Present
	<b>University of Illinois at Urbana-Champaign, Champaign, IL</b> <i>Postdoctoral Research Associate, with Prof. Heng Ji</i>	Sep 2021 – Aug 2022
	<b>Google AI, Mountain View, CA</b> <i>Research Intern</i>	May 2020 – Aug 2020
	<b>Allen Institute for Artificial Intelligence, Seattle, WA</b> <i>Research Intern</i>	Sep 2018 – Dec 2018
	<b>Microsoft Research, Redmond, WA</b> <i>Research Intern</i>	May 2018 – Aug 2018
SELECTED AWARDS& HONORS	<b>NSF CAREER award</b> <b>AAAI New Faculty Highlights</b> <b>Spotlight Rising Star in Data Science</b> <b>Top 100 New Stars in Artificial Intelligence</b> <b>Most Influential ACL Papers</b> (15 each year) <b>National Scholarship</b> (Top 3% undergrad students nationwide)	NSF, 2024 AAAI, 2024 University of Chicago, 2021 Baidu Scholar, 2020 Paperdigest, 2017 SJTU, 2013
PUBLICATIONS	<p>[1] <b>Leveraging Structured Information for Explainable Multi-hop Question Answering and Reasoning</b> Ruosen Li, <b>Xinya Du</b> <i>In Findings of the Association for Computational Linguistics: (EMNLP), 2023.</i></p> <p>[2] <b>Process of Elimination for Multiple Choice Reasoning</b> Chenkai Ma, <b>Xinya Du</b> <i>In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023.</i></p>	

- [3] **Probing Representations for Document-level Event Extraction**  
Barry Wang, **Xinya Du**, Claire Cardie  
*In Findings of the Association for Computational Linguistics: (EMNLP), 2023.*
- [4] **Zero-Shot Classification by Logical Reasoning on Natural Language Explanations**  
Chi Han, Hengzhi Pei, **Xinya Du**, Heng Ji  
*In Findings of the Association for Computational Linguistics: (ACL), 2023.*
- [5] **Toward Consistent and Informative Event-Event Temporal Relation Extraction**  
Xiaomeng Jin, Haoyang Wen, **Xinya Du**, Heng Ji  
*In MATCHING at Annual Meeting of the Association for Computational Linguistics (ACL), 2023.*
- [6] **End-to-end Case-Based Reasoning for Commonsense Knowledge Base Completion**  
Zonglin Yang, **Xinya Du**, Erik Cambria, Claire Cardie  
*In Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2023.*
- [7] **Logical Entity Representation in Knowledge-Graphs for Differentiable Rule Learning**  
Chi Han, Qizheng He, Charles Yu, **Xinya Du**, Hanghang Tong, Heng Ji  
*In International Conference on Learning Representations (ICLR), 2023.*
- [8] **Retrieval-Augmented Generative Question Answering for Event Argument Extraction**  
**Xinya Du** and Heng Ji  
*In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2022.*
- [9] **Dynamic Global Memory for Document-level Argument Extraction**  
**Xinya Du**, Sha Li, and Heng Ji  
*In Annual Meeting of the Association for Computational Linguistics (ACL), 2022.*
- [10] **Automatic Error Analysis for Document-level Information Extraction**  
Aliva Das, **Xinya Du**, Barry Wang, Kejian Shi, Jiayuan Gu, Thomas Porter, Claire Cardie  
*In Annual Meeting of the Association for Computational Linguistics (ACL), 2022.*
- [11] **RESIN-11: Schema-guided Event Prediction for 11 Newsworthy Scenarios**  
**Xinya Du**, Zixuan Zhang, Sha Li, Heng Ji and the RESIN team  
*In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL): System Demonstrations, 2022.*
- [12] **Template Filling with Generative Transformers**  
**Xinya Du**, Alexander M. Rush, and Claire Cardie  
*In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2021.*
- [13] **GRIT: Generative Role-filler Transformers for Document-level Event Entity Extraction**  
**Xinya Du**, Alexander M. Rush, and Claire Cardie  
*In Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2021.*

- [14] **Few-shot Intent Classification and Slot Filling with Retrieved Examples**  
Dian Yu, Luheng He, Yuan Zhang, **Xinya Du**, Panupong Pasupat and Qi Li  
*In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2021.*
- [15] **QA-Driven Zero-shot Slot Filling with Weak Supervision Pretraining**  
**Xinya Du**, Luheng He, Qi Li, Dian Yu, Panupong Pasupat and Yuan Zhang  
*In Annual Meeting of the Association for Computational Linguistics (ACL), 2021.*
- [16] **Event Extraction by Answering (Almost) Natural Questions**  
**Xinya Du** and Claire Cardie  
*In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020.*
- [17] **Improving Event Duration Prediction via Time-aware Pre-training**  
Zonglin Yang, **Xinya Du**, Alexander M. Rush and Claire Cardie  
*In Findings of the Association for Computational Linguistics: (EMNLP), 2020.*
- [18] **Document-Level Event Role Filler Extraction using Multi-Granularity Contextualized Encoding**  
**Xinya Du** and Claire Cardie  
*In Annual Meeting of the Association for Computational Linguistics (ACL), 2020.*
- [19] **Leveraging Structured Metadata for Improving Question Answering on the Web**  
**Xinya Du**, Adam Fourney, Robert Sim, Claire Cardie, Paul Bennett and Ahmed Hassan Awadallah  
*In Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AACL/IJCNLP), 2020.*
- [20] **Be Consistent! Improving Procedural Text Comprehension using Label Consistency**  
**Xinya Du**, Bhavana Dalvi, Niket Tandon, Antoine Bosselut, Wen-tau Yih, Peter Clark, Claire Cardie  
*In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2019.*
- [21] **Harvesting Paragraph-Level Question-Answer Pairs from Wikipedia**  
**Xinya Du** and Claire Cardie  
*In Annual Meeting of the Association for Computational Linguistics (ACL), 2018.*
- [22] **Identifying Where to Focus in Reading Comprehension for Neural Question Generation**  
**Xinya Du** and Claire Cardie  
*In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2017.*
- [23] **Learning to Ask: Neural Question Generation for Reading Comprehension**  
**Xinya Du**, Junru Shao and Claire Cardie  
*In Annual Meeting of the Association for Computational Linguistics (ACL), 2017.*  
**Featured in** *New Scientist* (“Inquisitive bot asks questions to test your understanding”), and *TechRepublic* (“How researchers trained one AI system to start asking its own questions”).
- [24] **Cornell Belief and Sentiment System at TAC 2016**  
Vlad Niculae, Kai Sun, Xilun Chen, Yao Cheng, **Xinya Du**, Esin Durmus, Arzoo Katiyar,

Claire Cardie  
*In Text Analysis Conference (TAC), 2016.*

TEACHING  
EXPERIENCE

Natural Language Processing, Cornell University, Fall 2019  
Teaching Assistant for Prof. Claire Cardie.

Natural Language Processing, Cornell University, Spring 2019  
Teaching Assistant for Prof. Yoav Artzi.

Software Engineering, Cornell University, Spring 17, Spring 18  
Teaching Assistant for Prof. William Arms.

Introduction to Computing Using Python, Cornell University, Fall 2016  
Teaching Assistant for Prof. Walker White.

MENTORING  
EXPERIENCE

Zonglin Yang, Cornell CS MEng student, 2020 – 2022.  
Topic: Commonsense and case-based reasoning for NLP.  
Publications: EMNLP (Findings), 2020.

Barry Wang, Cornell CS undergraduate student, 2021 – 2022  
Topic: Automatic error analysis for information extraction.  
Publications: ACL 2022, SciNLP 2021.

Aliva Das, Cornell CS undergraduate student, 2021 – 2022.  
Topic: Automatic error analysis for information extraction.  
Publications: ACL 2022, SciNLP 2021.

Maitreyi Chatterjee, Cornell CS undergraduate student, Spring 2021.  
Topic: Applying neural document-level IE model to scientific domain.

Rishi Malhotra, Cornell CS undergraduate student, Spring 2021.  
Topic: Applying neural document-level IE model to scientific domain.

PROFESSIONAL  
SERVICES

Chairing:  
ACL Rolling Review (ARR) 2024  
Annual Meeting of the Association for Computational Linguistics (ACL) 2023 (Area chair of Information Extraction Track)  
Natural Language Processing and Chinese Computing (NLPPCC) 2023 (Area chair of Question Answering Track)  
LREC-COLING 2024 (Area chair Information Extraction Track)  
Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2024 Website chair

Journal Reviewer:

IEEE Transactions on Knowledge and Data Engineering (TKDE)  
IEEE Transactions on Audio, Speech and Language Processing (TASLP)  
IEEE Transactions on Neural Networks and Learning Systems (TNNLS)  
ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)  
ACM Transactions on Knowledge Discovery from Data (TKDD)  
Knowledge and Information Systems (KAIS)  
AI Communication

Conference Committee Member:

Annual Meeting of the Association for Computational Linguistics (ACL)  
Conference on Empirical Methods in Natural Language Processing (EMNLP)  
Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)  
Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AACL/IJCNLP)  
International Joint Conference on Artificial Intelligence (IJCAI)  
AAAI Conference on Artificial Intelligence (AAAI)  
Natural Language Processing and Chinese Computing (NLPCC)  
Conference on Computational Natural Language Learning (CoNLL)  
Workshop on Noisy User-generated Text (W-NUT)  
Workshop on Machine Reading for Question Answering (MRQA)  
Joint Conference on Lexical and Computational Semantics

Volunteering and Other Activities:

Site host of North American Computational Linguistics Olympiad (NACLO) 2023.  
Member of Cornell CS Department PhD Admission Committee (Year 2021).  
Volunteer for Cornell CS Department PhD Visit Day (Year 2019, 2020, 2021).  
Student Volunteer for ACL 2017, ACL 2018, EMNLP 2017.

TALKS

**Towards More Informed Extraction of Events from Documents**

In Rising Stars in Data Science Workshop, University of Chicago, Jan 2021.

In Tencent AI Research, Nov 2020.

**Event Extraction by Answering (Almost) Natural Question**

In UIUC Class Information Extraction and Knowledge Acquisition, UIUC, Sep 2020

**LwLL: Progress on the NLP Front**

In DARPA site visit (Online), Cornell University, Apr 2020.

**Harvesting Paragraph-Level Question-Answer Pairs from Wikipedia**

In the 56th Annual Meeting of the Association for Computational Linguistics, July 2018.